Restaurant location recommender system in United States of America

Introduction:

Problem Background:

United States of America is fourth largest country in the world. Russia being the largest. Millions of Tourists come here to visit this beautiful country from all over the world. Asian food is one of the most liked dishes from around the world due to the famous spices put in it. People no matter where they belong from, they enjoy this food wherever they go to. Being of such an enormous size, it is difficult not to find the best resturant place nearby. So we shall focus on Larger cities only for this capstone project. USA has a population of 328 million with most of it being concentrated in major cities.

Problem Description:

People travel daily from their houses to work, travel up to 10 kms is considered normal in this city either by a vehicle or public transport. Travel takes times and hence food is the area which has be compromised upon. Hence the owner of ABC company has decided to start a recommender system for resturants in multiple cities based on their ratings available from foursquare. Asian food is a quick snack that can fill you up very fast. This recommender system can be used my many people around the country in famous cities. Just for this project I am concentrating on pizza but can be used for multiple things. So, there are a few things that should be considered.

1. Number of restaurants in that area.
2. The ones with good rating.
3. Number of similar restaurants (Asian restaurants) available near to me.
4. Average cost of the similar restaurants.

To help with this problem a recommender system is to get answer to the questions the manager of the ABC company is looking for.

Target Audience:

This is for people who travel a lot. People who work around the area and can come over during lunch break. This is also for people who are new to the area and looking for a quick snack. It is also for investors who are looking to invest in multiple areas and multiple properties. Finally, it is for everyone who is interested in trying something new.

Data

City data of the following will be selected:

- New York
- San Francisco
- Jersey City
- Boston
- Chicago

Latitude and Longitude coordinates of those neighborhoods. It is required to plot the data on the map.

Venue data about other restaurants taken from Foursquare.

Sources of Data

- We shall use foursquare data for all the cities and grab information from that.
- Then we shall get the coordinates of the same location using Python Geocoder Package.
- After that we shall use Foursquare API to get the venue location.

Methodology

My main target was to find the number of Asian restaurants available in Major cities of USA. I used Foursquare API to grab data. Data grabbed was of a pre categorized menu of Asian restaurants from the foursquare developer account.

4bf58dd8d48988d142941735 is the code for restaurants in major cities that fall under this category.

Further on I have repeated this request for an additional city's and got their top 100 locations that were available due to the restrictions on the foursquare developer account. These locations were plotted on to the map to visualize them.

Next was to find the center location of all the restaurants. A mean co= ordinate was found and using Euclidean distance between each restaurant from the mean point we find the mean distance. This can be used to find the average distance between the major hotels and the Asian restaurants in major cities.
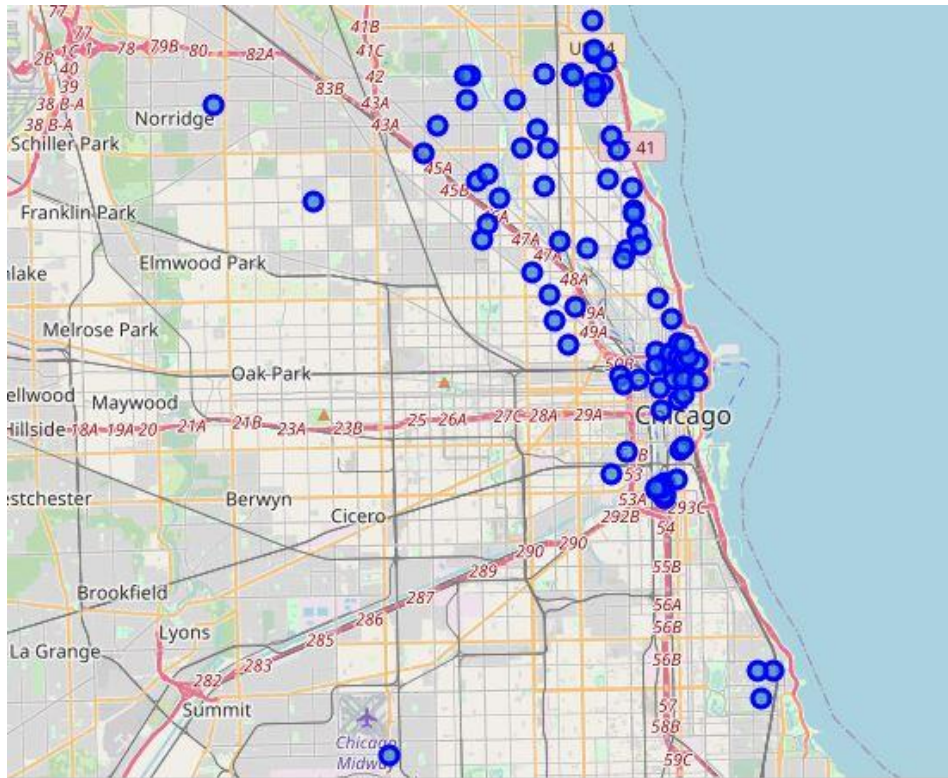
Results:

The following images of the map show that in which location the restaurants present.
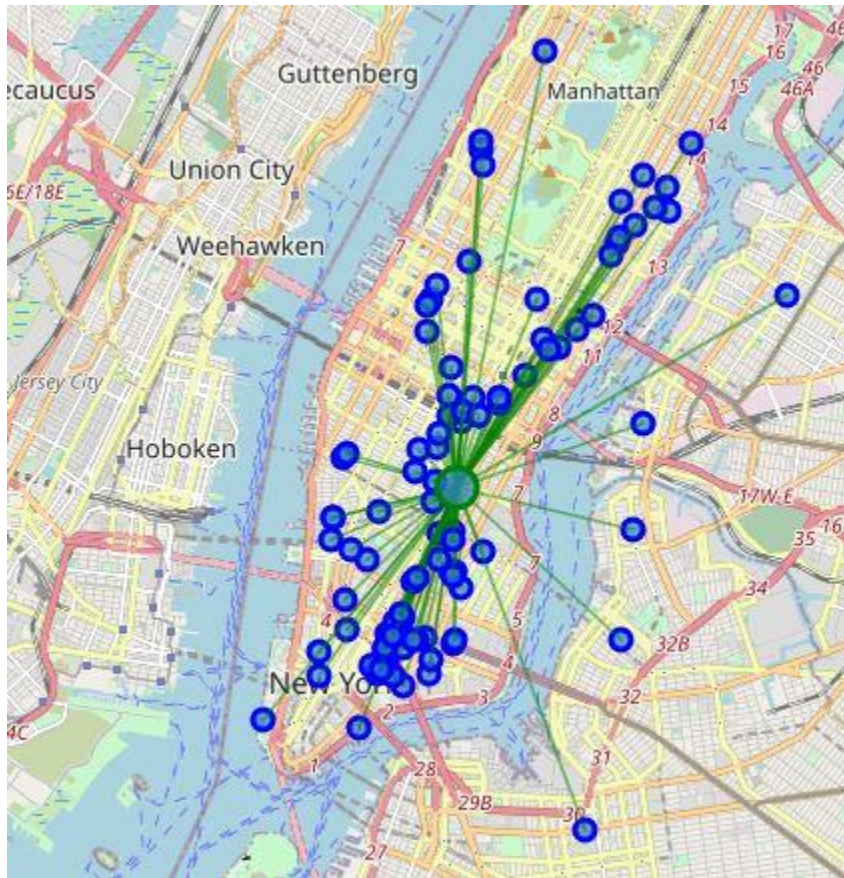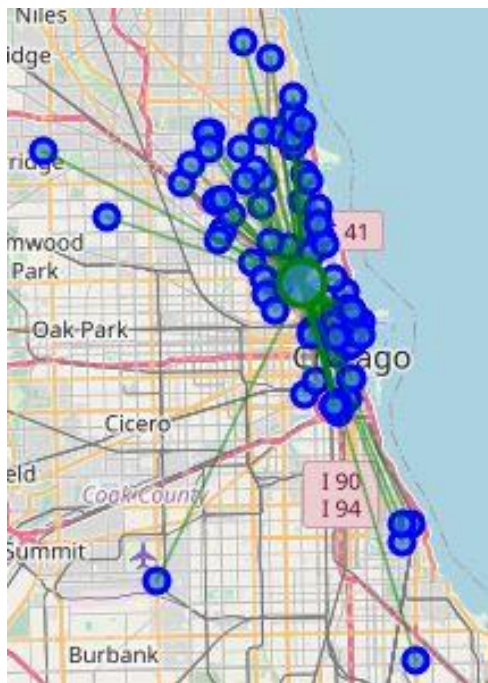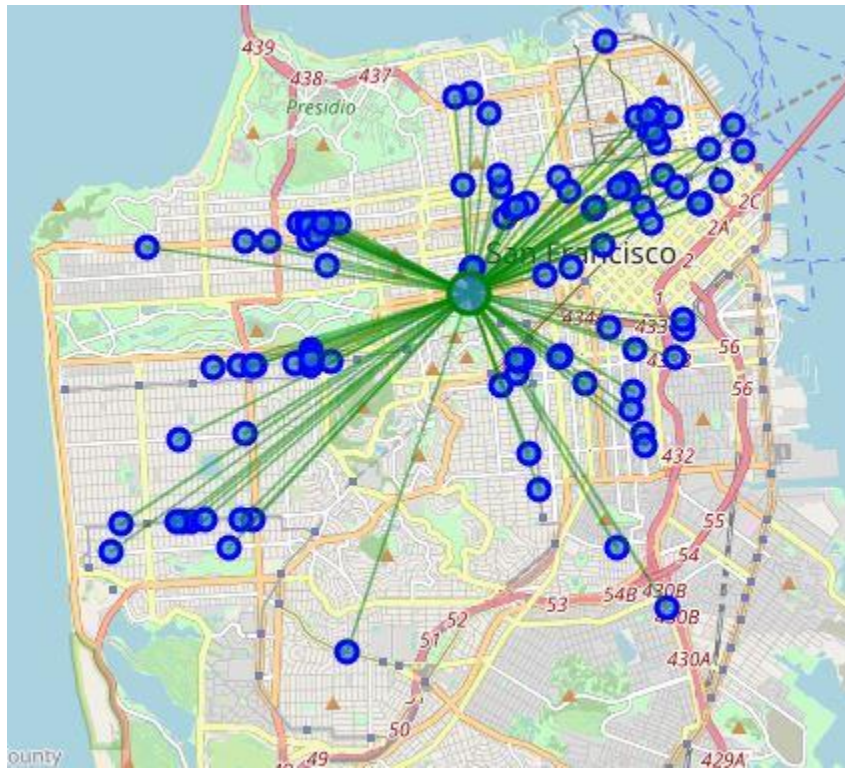
New York
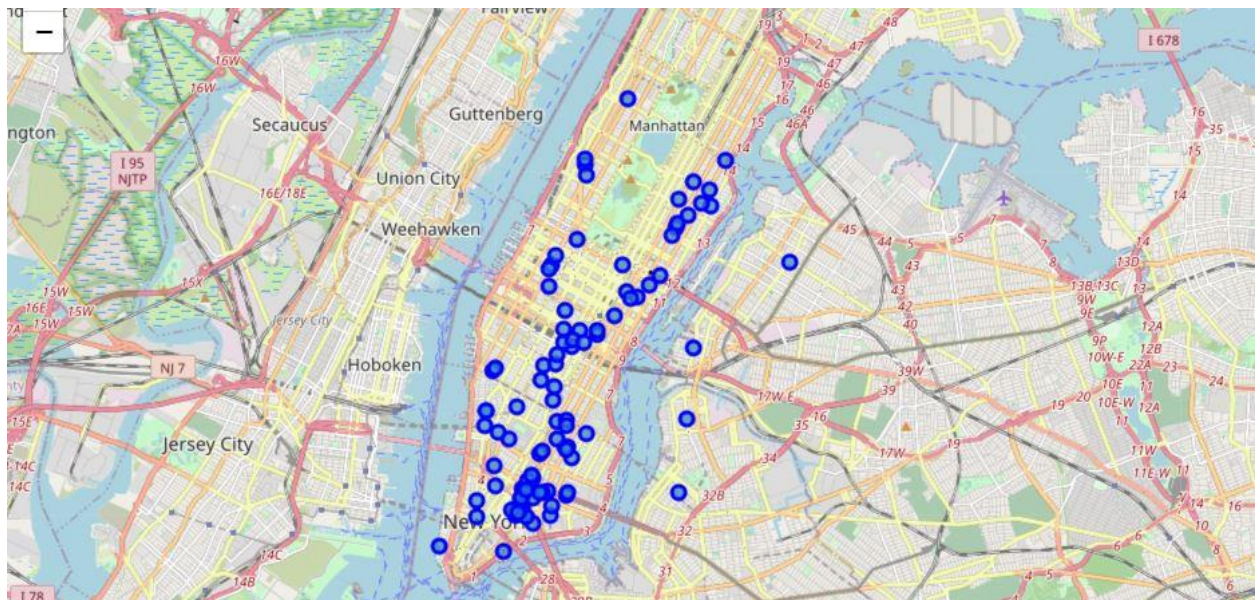
Chicago

San Francisco

New York

Chicago

San Francisco

Mean distances between properties.

```
New York, NY
Mean Distance from Mean coordinates
0.024946457390823514
Chicago, IL
Mean Distance from Mean coordinates
0.05590536362248579
San Francisco, CA
Mean Distance from Mean coordinates
0.03249653614237302
Jersey City, NJ
Mean Distance from Mean coordinates
0.01957299077845755
Boston, MA
Mean Distance from Mean coordinates
0.028305034100733546
```

The distance between Hotels and Asian Restaurants.

```python
city = 'Boston, MA'
venues_mean_coor = [df_venues[city]['Lat'].mean(), df_venues[city]['Lng'].mean()]

print(city)
print("Mean Distance from Mean coordinates")
dists = np.apply_along_axis(lambda x: np.linalg.norm(x - venues_mean_coor),1,df_venues[city][['Lat','Lng']].values)
dists.sort()
print(np.mean(dists[:-1]))# Ignore the biggest distance
```
```
Boston, MA
Mean Distance from Mean coordinates
0.027619831836190874
```

```python
city = 'New York, NY'
venues_mean_coor = [df_venues[city]['Lat'].mean(), df_venues[city]['Lng'].mean()]

print(city)
print("Mean Distance from Mean coordinates")
dists = np.apply_along_axis(lambda x: np.linalg.norm(x - venues_mean_coor),1,df_venues[city][['Lat','Lng']].values)
dists.sort()
print(np.mean(dists[:-1]))# Ignore the biggest distance
```
```
New York, NY
Mean Distance from Mean coordinates
0.02457306207948573
```

Conclusion:

Hence, we can say that we have found the number of restaurants in major cities. The Highest to lowest is as follows.

1. New York – 332
2. Chicago – 304
3. San Francisco – 301
4. Boston – 270
5. Jersey City – 167

But since only 100 information is shared by foursquare. We can only show that and not the rest.

We have also found the mean distance from the center point to the farthest point. We have also found the mean distance to hotels so as to tourist can also enjoy the restaurants.